Appendix B.
Sample Design and Weighting

SAMPLE SIZE

The 2011 national data are from a sample of housing units interviewed between July and December 2011. The same basic sample of housing units is interviewed every two years until a new sample is selected. The U.S. Census Bureau updates the sample by adding newly constructed housing units and units discovered through coverage improvement efforts.

For the 2011 American Housing Survey-National (AHS-N) sample, approximately 56,200 sample housing units were originally selected for interview. The Department of Housing and Urban Development (HUD) requested that the AHS-N sample size be increased to 64,750. Therefore, two types of samples were reinstated in 2011. First, the roughly 5,200 units that were reduced from the 2007 sample[1] were reinstated. Second, approximately 3,100 units that were selected during the 2000 redesign, but were never interviewed, were introduced in 2011. Later, HUD requested an oversample of approximately 5,300 subsidized housing units. Lastly, the American Housing Survey-Metropolitan Sample (AHS-MS) was combined with the AHS-N for the first time in the survey's history, adding approximately 116,700 housing units to the national sample. These twenty-nine metropolitan areas from AHS-MS were used as a supplemental sample. Sample sizes for these metropolitan areas are included in Table B-1. Therefore, the sample size for the 2011 national sample was approximately 186,400 housing units.

About 8,900 of the 186,400 total units included for interview were found to be ineligible because the unit no longer existed or because the units did not meet the AHS-N definition of a housing unit.

Of the 177,500 eligible sample units, about 22,800 were classified (both occupied and vacant housing units) as "Type A" noninterviews because (a) no one was at home after repeated visits, (b) the respondent refused to be interviewed, or (c) the interviewer was unable to find the unit.
This classification produced an unweighted overall response rate of 87 percent. The weighted overall response rate was 88 percent.

SAMPLE SELECTION

The Census Bureau has interviewed the current sample of housing units since 1985. First, the United States was divided into areas made up of counties or groups of counties and independent cities known as primary sampling units (PSUs). A sample of these PSUs was selected. Then, a sample of housing units was selected within these PSUs.

Selection of sample areas. The sample for AHS is spread over 394 PSUs. These PSUs cover 878 counties and independent cities with coverage in all 50 states and the District of Columbia.

If there were over 100,000 housing units in a PSU at the time of selection, the PSU is known as a self-representing PSU because it was removed from the probability sampling operation. It was in sample with certainty. The sample from the PSU represents only that PSU. There are 170 self-representing PSUs.

The Census Bureau grouped the remaining PSUs into similar-sized strata, based on similarities in various socioeconomic characteristics within each stratum, and selected one PSU per stratum, proportional to the number of housing units in the PSU, to represent all PSUs in the stratum. These selected PSUs are referred to as non-self-representing PSUs. The sample non-self-representing PSUs for AHS are a subsample of the Current Population Survey's (CPS) sample areas based on the 1980 census.

Selection of sample housing units. The AHS sample consists of the following types of units in the sampled PSUs:

- Housing units selected from the 1980 census
- New construction in areas requiring building permits
- Housing units missed in the 1980 census
- Other housing units added since the 1980 census
- Housing units selected from the 2000 census
- Subsidized housing units

---

[1] Due to budgetary constraints, roughly 8 percent of units were taken out of the 2007 sample and were not interviewed until 2011.

Housing units selected from the 1980 census. The Census Bureau picked a systematic sample so every unit had a 1 in 2,148 chance of being included in the AHS.

In areas where addresses are complete (at least 96 percent of units having a house number and street name) and permits are required for new construction, housing units receiving 1980 census long-form questionnaires were sorted by the following items:

- PSU
- Central city, urbanized area, urban outside urbanized area, rural
- Owner, renter, vacant for rent, vacant for sale, other types of vacant housing units
- Number of rooms
- Value of housing unit or gross rent
- Manufactured/mobile home or not a manufactured/mobile home

In areas where addresses are not complete or permits are not required for new construction, land areas were sorted using a formula incorporating the following items:

- PSU
- Central city, urbanized area, urban outside urbanized area, rural
- Median value of housing unit
- Number of children under 6 years old
- Number of elderly people
- Number of owner-occupied housing units
- Number of manufactured/mobile homes
- Number of housing units lacking some plumbing
- Number of owner-occupied housing units whose value is below $45,000
- Number of renter-occupied housing units with rent below $200
- Number of Black and Hispanic people
- Number of 1-room housing units

New construction in areas requiring building permits. In areas that require building permits for new construction, the Census Bureau selected a sample of permits. These permits do not cover manufactured/mobile homes or conversion of buildings to residential use.

Housing units missed in the 1980 census. The Census Bureau conducted a special study which identified units at addresses missed or inadequately defined in the 1980 census. A sample of these identified units was selected.

Other housing units added since the 1980 census. If extra units are added in buildings or manufactured/mobile home parks where AHS already has sample units, a sample of these extra units was selected. To find when whole buildings are built (in addition to building permits, mentioned above) or are converted from nonresidential to residential use, the Census Bureau listed all residential buildings in a sample of areas around the country, found any additional buildings, and selected a sample of their units.

Housing Units Selected From Census 2000
The following adjustments were made to the AHS-N sample in 2005 by adding certain types of units selected from Census 2000:

In 2005, a new sample of manufactured/mobile homes was selected from Census 2000 in an attempt to improve coverage of manufactured/mobile homes built between 1980 and 2000. These units were selected at the same rate as the 1980-based units. One-half of this sample was included in the 2005 interviewing and, as a result, one-half of the 1980-based sample was not included. The units interviewed in 2005 were also interviewed in 2007, 2009, and 2011.

In 2005, a sample of assisted living units was selected from Census 2000 in an attempt to improve coverage of the elderly. The Census Bureau purchased address lists of assisted living facilities from two vendors and supplemented these lists with facility addresses from various assisted living facility internet sites. These addresses were matched to Census 2000 to identify Census 2000 housing units in assisted living facilities from which the resulting sample of these units was drawn. Because the universe was so small, a double sample of units was selected. Before interviewing, the resulting sample was screened to confirm eligibility. There were 921 units in the initial sample with 486 remaining after screening. The units interviewed in 2005 were also interviewed in 2007, 2009, and 2011.

Subsidized Housing Units
A sample of subsidized housing units was selected from address lists provided by HUD in an attempt to improve coverage of housing units receiving rent subsidies. The Census Bureau received address lists from HUD in 2010. These lists include the Public and Indian Housing Information Center (PIC), Tenant Rental Assistance Certification System (TRACS), and Home Investment Partnership Program (HOME) address lists. The lists were matched to the sample PSUs to enable the Census Bureau to select housing units receiving HUD subsidies. There were 5,259 units in the initial sample with 5,064 eligible for interview.

SUPPLEMENTAL METROPOLITAN SAMPLE

In 2011, the Census Bureau supplemented the national sample in 29 metropolitan areas. This supplemental sample was combined with the existing national sample in these areas in order to produce metropolitan level estimates and to increase the sample size for the AHS-N sample. The housing units that were part of the AHS-N sample are based on the definitions used in 1985, as was the Los Angeles AHS-MS sample. The Providence AHS-MS sample was based on the 2003 Office of Management and Budget (OMB) definition for the New England City and Town Area Division (NECTAD). The remaining AHS-MS sample is consistent with the 2003 OMB definitions of the metropolitan statistical area (MSA), with the following exceptions:

- The Anaheim MSA matches the December 2003 OMB definition for the Santa Ana-Anaheim-Irvine Metropolitan Division.

- The Dallas MSA matches the 2003 OMB definition for the Dallas-Plano-Irving Metropolitan Division.

- The Denver MSA does not include Broomfield County. The county borders that existed before Broomfield's creation were used in the AHS definition.

- The Fort Worth MSA matches the 2003 OMB definition for the Fort Worth-Arlington Metropolitan Division.

- The Los Angeles MSA matches the 2003 OMB definition for the Los Angeles-Long Beach-Glendale Metropolitan Division.

- The Oakland MSA matches the 2003 OMB definition for the Oakland-Fremont-Hayward Metropolitan Division.

- Included in the Providence NECTAD but not in the 2003 OMB definition for the Providence-New Bedford-Fall River MSA are: Bellingham and Plainville (in Norfolk County, MA); Blackstone and Millville (in Worcester County, MA). Included in OMB's 2003 definition of the Providence-New Bedford-Fall River MSA but not in the Providence MSA are: Acushnet, Berkley, Dartmouth, Dighton, Easton, Fairhaven, Freetown, Mansfield, New Bedford, Norton, Raynham, and Taunton (these are in Bristol County, MA); and New Shoreham and Westerly town (in Washington County, RI).

- The San Francisco MSA matches the 2003 OMB definition for the San Francisco-San Mateo-Redwood City Metropolitan Division.

- The St. Louis MSA does not include Sullivan City, which is legally part of OMB's 2003 definition, though it resides in an outlying county.

In some areas, the following adjustments were made:

- Counties/Minor Civil Divisions (MCDs) were added or dropped so that the definition of each metropolitan area in sample was consistent with the final 2003 OMB definition of the metropolitan area and sample was selected in these added areas.

- The sample in the counties/MCDs in the previous definition that were also in these new definitions (i.e., continuing counties/MCDs) was adjusted to maintain an overall sample size of 4,500 and in some cases it was replaced by new sample for confidentiality reasons.

Table B-1 provides the size of the supplemental sample in each of the 29 metropolitan areas. This sample was combined with the existing sample in these areas to produce metropolitan estimates.

Table B-1.

2011 Sample Size for the 29 AHS-National-Based Metropolitan Areas (in housing units)

Metropolitan Area	Basic Sample	Supplemental Sample	Total Sample Size
Anaheim, CA……………	517	4,011	4,528
Atlanta, GA………………	991	3,578	4,569
Birmingham, AL………..	296	4,387	4,683
Buffalo, NY……………..	308	4,149	4,457
Cincinnati, OH…………..	434	4,132	4,566
Cleveland, OH…………..	551	4,129	4,680
Columbus, OH…………..	427	4,157	4,584
Dallas, TX………………..	787	3,827	4,614
Denver, CO………………	673	3,777	4,450
Fort Worth, TX………….	401	4,184	4,585
Indianapolis, IN…………	415	4,144	4,559
Kansas City, MO………..	581	3,978	4,559
Los Angeles, CA………..	1,884	2,708	4,592
Memphis, TN…………….	272	4,233	4,505
Milwaukee, WI………….	390	4,203	4,593
New Orleans, LA………..	301	4,545	4,846
Virginia Beach, VA……..	403	4,249	4,652
Phoenix, AZ……………..	808	3,731	4,539
Pittsburgh, PA………….	652	3,955	4,607
Portland, OR…………….	608	4,019	4,627
Providence, RI………….	316	4,368	4,684
Riverside, CA……………	735	3,902	4,637
San Diego, CA………….	595	3,967	4,562
San Francisco, CA………	430	4,085	4,515
San Jose, CA……………	339	4,153	4,492
St. Louis, MO……………	652	3,917	4,569
Charlotte, NC……………	522	4,100	4,622
Oakland, CA…………….	555	3,995	4,550
Sacramento, CA…………	462	4,118	4,580

Table B-2.

Interview Activity by AHS Sample

Metropolitan area	Unweighted response rate[2] (percent)	Weighted response rate[3] (percent)	Eligible units			Ineligible[5]
			Total	Inter-viewed	Not inter-viewed[4]	
TOTAL..…………………	87.2	87.6	177,506	154,752	22,754	8,892
AHS-MS.…………………	87.1	88.0	110,366	96,092	14,274	6,335
AHS-N…..………………	87.4	87.5	67,140	58,660	8,480	2,557

Table B-2 summarizes the interview activity for the two AHS samples, and for the combined sample. The table provides the response rate, number of eligible units (comprised of completed interviews and noninterviews), and the number of units visited but ineligible for interview.

ESTIMATION FOR AHS-NATIONAL

Each housing unit in the AHS sample represents itself and over 2,000 other units. The exact number it represents is its "weight." The weight was calculated in six steps. The purpose of these steps is to minimize both sampling errors and errors from incomplete data. The result of these steps is also to force consistency with some major categories of data in other Census Bureau surveys, so figures on these categories do not actually depend on the AHS sample, but on the other surveys.

For the first time in the survey's history, Census division-level estimates were calculated. Therefore, the adjustments made to force consistency with other Census Bureau surveys were done so for each of the nine Census divisions.

1. **Basic weight.** The Census Bureau assigned each unit a weight to reflect its probability of selection. With rare exceptions, the AHS-National weight is 2,148. One exception is the special living sample which is assigned a weight of 1,074. In 2011, the AHS-Metropolitan Sample was combined with the AHS-National Sample. Since each metropolitan area was sampled independently, each metropolitan area had its own sampling rate and therefore had its own basic weight. These weights ranged approximately between 100 and 600.

2. **Sample adjustment.** An adjustment was made to the units to account for the introduction of the 2000-design cases, the addition of the supplemental sample in the 29 metro areas, and for the addition of the oversample of subsidized housing units. This adjustment was made to ensure the additional sample would not inflate the national housing unit estimates.

3. **Noninterview adjustment.** An adjustment was made for refusals and occupied units where no one was home. The calculations for this adjustment do not include units the Census Bureau could not locate. The earlier weight was multiplied by the following factor:

---

[2] The unweighted response rate is computed by dividing the unweighted number of interviews by the unweighted total number of cases eligible for interview and multiplying by 100.

[3] The weighted response rate is computed by dividing the weighted number of interviews by the weighted total number of cases eligible for interview and multiplying by 100.

[4] Sample units were classified as "Type A" noninterviews because (a) no one was at home after repeated visits, (b) the respondent refused to be interviewed, or (c) the interviewer was unable to find the unit.

[5] Sample units were found to be ineligible because the unit no longer existed or because the unit did not meet the AHS definition of a housing unit.

$$\frac{\text{Interviewed units + Units not interviewed}}{\text{Interviewed units}}$$

It is assumed the units missed are similar in some ways to the units interviewed for AHS.

This adjustment is done separately for groups defined by cross-classifying the following data items if prior year data for the indicated items are available:

- Twenty-nine AHS-MS metropolitan areas
- Nine Census divisions (for units that are not in an AHS-MS metropolitan area)
- 1990 central city, suburb, or nonmetropolitan area
- 1990 urban or rural
- Manufactured/mobile home or not a manufactured/mobile home
- In a special living facility or not in a special living facility (only if the housing unit is not a manufactured/mobile home)
- Owner/for sale or renter/for rent
- Number of units in structure*
- Number of rooms*
- Occupied, vacant year round, or seasonal/migratory vacant*

(* If known from a previous survey; otherwise, the Census Bureau substituted whether or not units were drawn from building permits for these items.)

For seasonal/migratory vacants and year round vacants other than those for rent or for sale, units were cross-classified only by census region and 2000 central city/suburb/nonmetropolitan area.

4. **PSU adjustment.** The Census Bureau adjusted for differences that existed in 1990 between the number of 1990 census housing units estimated from the AHS sample of non-self-representing PSUs and the 1990 census counts outside the self-representing PSUs. The earlier weight was multiplied by the following factor:

$$\frac{\text{1990 census housing units in all areas that}}{\text{could have been chosen as non-self-representing PSUs}}$$
$$\text{1990 census housing units estimated from the}$$
$$\text{AHS sample of non-self-representing PSUs}$$

This adjustment is done separately for groups defined by cross-classifying:

- Four Census regions
- Owner, renter, or vacant
- 1990 central city, suburb, or nonmetropolitan area
- 1990 urban or rural
- Hispanic or non-Hispanic householder (only in South and West regions)
- Black or non-Black householder (only in South region)

5. **New construction adjustment.** The Census Bureau adjusted for known deficiencies in sampling new construction by multiplying the earlier weight by the following factor:

$$\frac{\text{Independent estimate}}{\text{AHS sample estimate}}$$

This adjustment is done separately for groups defined by cross-classifying:

- Nine Census divisions
- Manufactured/mobile home or not a manufactured/mobile home
- Number of units in structure
- Year built (pre-1980 and five-year categories after 1980 as shown in the publication)

Independent estimates are based on the Census Bureau's Survey of Construction and Survey of Manufactured Home Placements. Note that final AHS figures for the categories above are not based on the AHS sample findings, but on the independent sources.

6. **Demographic adjustment.** Comparability among the surveys was ensured by multiplying the earlier weight by the following factor:

$$\frac{\text{Independent estimate}}{\text{AHS sample estimate}}$$

This adjustment is done in two steps for occupied units. First, the factors were computed and applied for the Hispanic or non-Hispanic groups defined by cross-classifying:

- Nine Census divisions
- Owner or renter
- Hispanic or non-Hispanic householder
- Husband-wife, other male householder, or other female householder
- Age of householder

Next, the demographic adjustment is repeated with the same cells, except classified by the Black or non-Black groups, rather than the Hispanic or non-Hispanic groups.

Vacant for sale, vacant for rent, other year round vacant, and seasonal/migratory vacant units were cross-classified only by the four census regions and 2000 central city, suburb, or nonmetropolitan area.

The percentage of occupied and vacant units was based on the AHS itself. The distributions within occupied and vacant units are from the Census Bureau's Current Population Survey for occupied units and from the Housing Vacancy Survey for vacant units. The total number of all housing units in the United States is based on the 2010 Census adjusted to account for new and lost units. Note that final AHS figures for the categories above are not based on the AHS sample findings, but on the independent sources.

Repetitions. The new construction and demographic adjustments were repeated to help match both sets of independent estimates simultaneously. These adjustments were repeated until every cell's factor is between 0.98 and 1.02 or the change in each factor from one repetition to the next is less than 0.015.

Small cells. In each step of weighting, many items were cross-classified; so some cells may have few cases. When a cell is too small (less than 30 cases for the noninterview adjustment or less than 50 cases for the demographic adjustment) or the adjustment factor is too extreme (greater than 1.5 for the noninterview adjustment or outside a range of 0.5 to 2.0 for the demographic adjustment), the Census Bureau combined the cell with one or more other cells that are similar in most respects. Cells for the PSU adjustment or the new construction adjustment were not combined.